Truck Weight Distributions at Traffic Count Sites Using WIM and GPS Data

Sarah Hernandez, PhD
Assistant Professor
Department of Civil Engineering
University of Arkansas
4190 Bell Engineering Center
Fayetteville, Arkansas 72701
sarahvh@uark.edu
Office: 479-575-4182
Fax: 479-575-7168

Kyung (Kate) Hyun
PhD Candidate
Department of Civil Engineering
University of California, Irvine
4000 Anteater Research and Instruction Building
Irvine, California 92697
khyun1@uci.edu
Office: 949-824-5989
Fax: 949-824-8385

Words: 5,248 words
Tables: 1 x 250 = 250 words
Figures: 8 x 250 = 2,000 words
Total: 7,498 words

ABSTRACT

This paper presents a method for estimating gross vehicle weight (GVW) distributions of five axle tractor trailers ('3-S2') at traffic count sites. Traffic count sites include inductive loop detector or automatic vehicle classifier sites that do not measure vehicle weight unlike Weigh-in-Motion (WIM) sites. Truck weight data is needed for pavement design, weight enforcement, traffic monitoring, and freight transportation planning but the low spatial resolution of WIM can limit potential applications. This paper proposes to increase the spatial resolution of truck weight data by providing weight estimates at traffic count sites using Gaussian Mixture Models (GMM). A GVW distribution at a traffic count site is estimated by combining GMMs estimated at WIM sites that are defined to be spatially related to the traffic count site. Truck travel patterns derived from a large truck GPS database are used to determine the degree to which a WIM and traffic count site are spatially related. Specifically, the number of GPS truck traces that cross both the traffic count site and each WIM site defined the mixing proportions in the GMM. A leave-one-out cross validation framework allows for comparisons of estimated and measured GVW distributions at each WIM site. Coincidence ratios and two-sample Kolmogorov-Smirnov (KS) tests are used as comparison metrics for a case study of 112 WIM sites in California. The proposed methodology provides favorable results compared to a baseline approach which defines the spatial relation between sites using Great Circle Distances (GCD).

1 INTRODUCTION

Truck weight data is a key input for pavement design and performance monitoring, weight enforcement, bridge performance, freight transportation planning, and emissions estimation. In pavement design, under the Mechanical Empirical Pavement Design Guide (MEPDG) truck weight data is needed at a high level of spatial resolution. In fact, 'Level 1' analysis in the MEPDG requires local truck count and weight data. Because localized truck weights can be cost prohibitive to obtain, it is common to use state or national averages in place of localized data. This can lead to inefficient pavement designs such as unnecessarily thick or inadequately thin asphalt pavement layers, for example (*1, 2*). Freight forecasting models use truck weight data to define average payloads used to convert predictions of commodity tonnages to numbers of truck trips. The Freight Analysis Framework (FAF) derives average payloads solely from surveys (*3*). However, these surveys may not provide representative samples of state-level truck characteristics like average payload and loaded weights due to the structure of the survey (*4*). Emissions estimation tools use vehicle classification schemes based on truck weights to apply emissions rates (*5*). Vehicle registration databases and truck surveys are used to supply input data for the models but it may be beneficial to use measured weights taken by field data collection devices. Truck weight data has also been used to assess the impact of truck loads on the performance and durability of bridges (*6*).

Despite the need for comprehensive truck weight data, the current methods for collecting such data provide coverage for only a small proportion of a state's roadway network. In California, for example, there are around 110 weight-capable traffic data collection devices, i.e. Weigh-In-Motion (WIM) sites, providing weight data for 174,991 miles of maintained roadway (*7*). The limited coverage is due in part to the expense of installing and maintaining WIM which can cost as much as $50,000 per lane for installation and $8,000 per year for maintenance and operation (*6*). Typically located on major truck routes, WIM capture truck axle configurations and weights at mainline highway speeds in order to provide measurements of gross vehicle weight (GVW), speed, axle weight, and axle spacing (*8*). Static weight stations, or enforcement stations, are limited by even sparser coverage and data from these sites are not typically archived for historical analysis.

Less costly and more widely deployed traffic count devices such as inductive loop detectors (ILDs) or automatic vehicle classifiers (AVC) are capable of collecting traffic counts and in some configurations provide traffic counts by vehicle type (*9*). As a reference, California maintains over 8,000 ILD sites in their Performance Measurement System (PeMS) (*10*). However, ILDs and other available temporary and permanent traffic count devices such as tube counters, magnetometers, radar, and LIDAR do not measure truck weight. New and emerging sources of truck data including mobile sensors like Global Positioning Systems (GPS), smartphone sensors, and connected vehicles systems, while valuable sources of travel time and origin-destination data, likewise do not provide information about vehicle weights.

Considering the many applications of comprehensive weight data and given the limited availability of such data, this paper applies mathematical modeling to provide estimates of truck GVW distributions at traffic count sites. Each traffic count site that currently only collects traffic counts can be used to provide localized truck GVW distributions to suit many of the applications previous described without requiring additional data collection.

2 METHODOLOGY

The methodology entails combining GVW distributions measured at WIM sites based on truck routing patterns gathered from truck GPS data. The methodology outlined in **Figure 1** can be applied to five axle semi- tractor trailers, also known as '3S2' or FHWA Class 9 trucks, the most common configuration for freight trucks. After applying a normalization process to ensure the weight data is not skewed as a result of an uncalibrated weight sensor (**Step 1**), a GMM is fit to the GVW distribution at each WIM site (**Step 2**). The best fit GMM consists of either two or three components which are represented as individual Gaussian distributions. Using truck GPS data, the number of truck trips passing between a traffic count site and each WIM in the sensor network are used to estimate a Spatial Relation Matrix (**Step 3**). The Spatial Weight Matrix defines the mixing proportions that are applied to the GMM components of each WIM site (**Step 4**). Specifically, the number of GPS truck traces that cross both the traffic count site and each WIM site defined spatial weighting measures. Finally, the GVW distribution at the traffic count site is determined by combining each WIM site's GMM components using the mixing proportions defined by the Spatial Weight Matrix (**Step 5**).

The final output of the methodology is a frequency distribution of GVW. To get a distribution of GVW (i.e. histogram), the GVW frequency distribution can be multiplied by the estimated count of '3-S2' trucks at the traffic count site. If the traffic count site provides vehicle classification counts such as at an AVC site, then the GVW frequency distribution is multiplied by the class count to give a distribution of GVW for the class. For ILD sites, several promising methods have been introduced to estimate truck classification using advanced ILDs (*11, 12*).

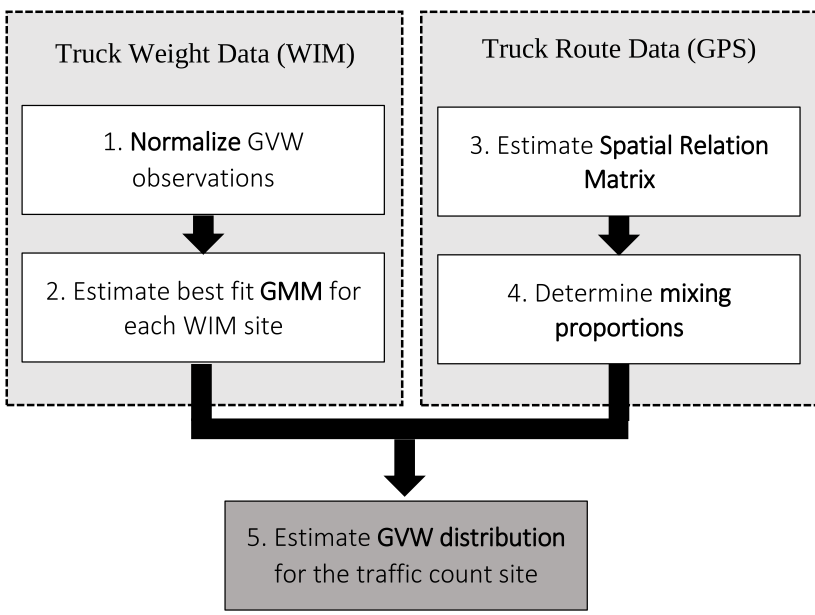


FIGURE 1 Overview of methodology.

2.1 Normalization of GVW Weight Distributions

WIM sensors can contain both random and systematic measurement errors as a result of vehicle dynamics over the sensor, site conditions, and environmental factors (*13, 14, 15*). Systematic errors are persistent inaccuracies which exhibit as over- or under- estimations of the true weights. Methods have been developed to assess calibration issues (*16, 17*) and to correct weight data that is was collected from an out-of-calibration sensor (*18, 19, 20*). Random errors are more difficult to detect and no correction procedures have been introduced in the literature to date.

For this paper, a simple approach to correct for possible systematic weight errors was employed. Since the steering axles of '3-S2' trucks have relatively constant weights of 10 kips across locations (*20*), a station reporting steering axle weights systematically above or below 10kips is likely out of calibration. In this paper, GVW was normalized by dividing by steering axle weight as follows:

$$GVW_t^{norm} = \frac{GVW_t}{AX1_t}$$

where
GVW_t^{norm} is the normalized GVW of truck t (unitless)
GVW_t is the measured GVW of truck t in kips recorded by the WIM sensor
$AX1_t$ is the steering axle weight in kips

Denormalization of the final estimated GVW distribution requires multiplying the normalized GMM components (e.g. means and covariance) by an assumed steering axle weight of 10 kips. In future work, more complex weight data correction schemes can be easily integrated into this stage of the methodology.

2.2 Gaussian Mixture Models (GMM)

Once the weight data has been normalized to remove effects of sensor calibration issues, a GMM of the GVW distribution was estimated for each WIM site. A GMM is a linear composition of individual Gaussian distributions, $\mathcal{N}(x; \mu_{i,m}, \Sigma_{i,m})$, combined via a mixing parameter, $p_{i,m}$, as follows (*21*):

$$f_i(x) = \sum_{m=1}^{M} p_{i,m} \cdot \mathcal{N}(x; \mu_{i,m}, \Sigma_{i,m})$$

where
$f_i(x)$ is the distribution of GVW's (x) for site i
$\mathcal{N}(\mu_{i,m}, \Sigma_{i,m})$ is the Gaussian distribution with mean μ_m and covariance matrix Σ_m, m = 1…M for site i where M is the maximum number of components in the GMM
$p_{i,m}$ is the mixing proportion for the m^th GMM component at site i such that $\sum_{m=1}^{M} p_{i,m} = 1$

To estimate a GMM, the number of component distributions, M, must be predetermined. Previous research has shown that for '3-S2' trucks the GVW distribution can be represented as a two or three component GMM (*22*). With three components, the first component is assumed to represent the weight distribution of empty trucks, the second represents partially loaded trucks, and the third represents fully loaded trucks. For mixtures of two components, the first component represents the distribution of empty weight while the second represents the distribution of loaded trucks. Interestingly, Hyun et al. (*23*) showed that the number of components is related to the body type of the trailer, i.e. tank, van, flatbed. Tanks which carry liquid commodities travel either empty or loaded while vans tend to have tri-modal distributions. For each WIM site, a two and a three component GMMs were fit to the GVW distribution. The best-fit GMM for each WIM site was then chosen based on the Akaike Information Criteria (AIC) where the smaller value of the AIC corresponds to the best fit model (*24*).

2.3 *Spatial Relation Matrix and Mixing Proportions*

The underlying theory for the proposed methodology is that a weight distribution at a traffic count site can be estimated by combining the weight distributions from WIM sites that see the same truck traffic as the traffic count site. There is both a spatial range and directionality to the GVW distribution patterns. Sensors along the same route in the same direction and within the same region share similar GVW distribution patterns. For example, **Figure 2** compares the GVW distributions of five axle semi-tractor trailers traveling in the north- and southbound directions across four WIM sites in northern California. Each of the sites has a significant volume of loaded trucks as evidenced by the high peak in the upper weight range. These four sites see a large proportion of the same truck trips due to commodity flow patterns within the northern California region. Differences in weight distributions can also be observed in the directional flows where the northbound (NB) and southbound (SB) directions at a WIM site may exhibit opposite GVW patterns. Similarly, WIM sites along parallel-routes, although within the same region and in the same direction may not have similar GVW distributions.

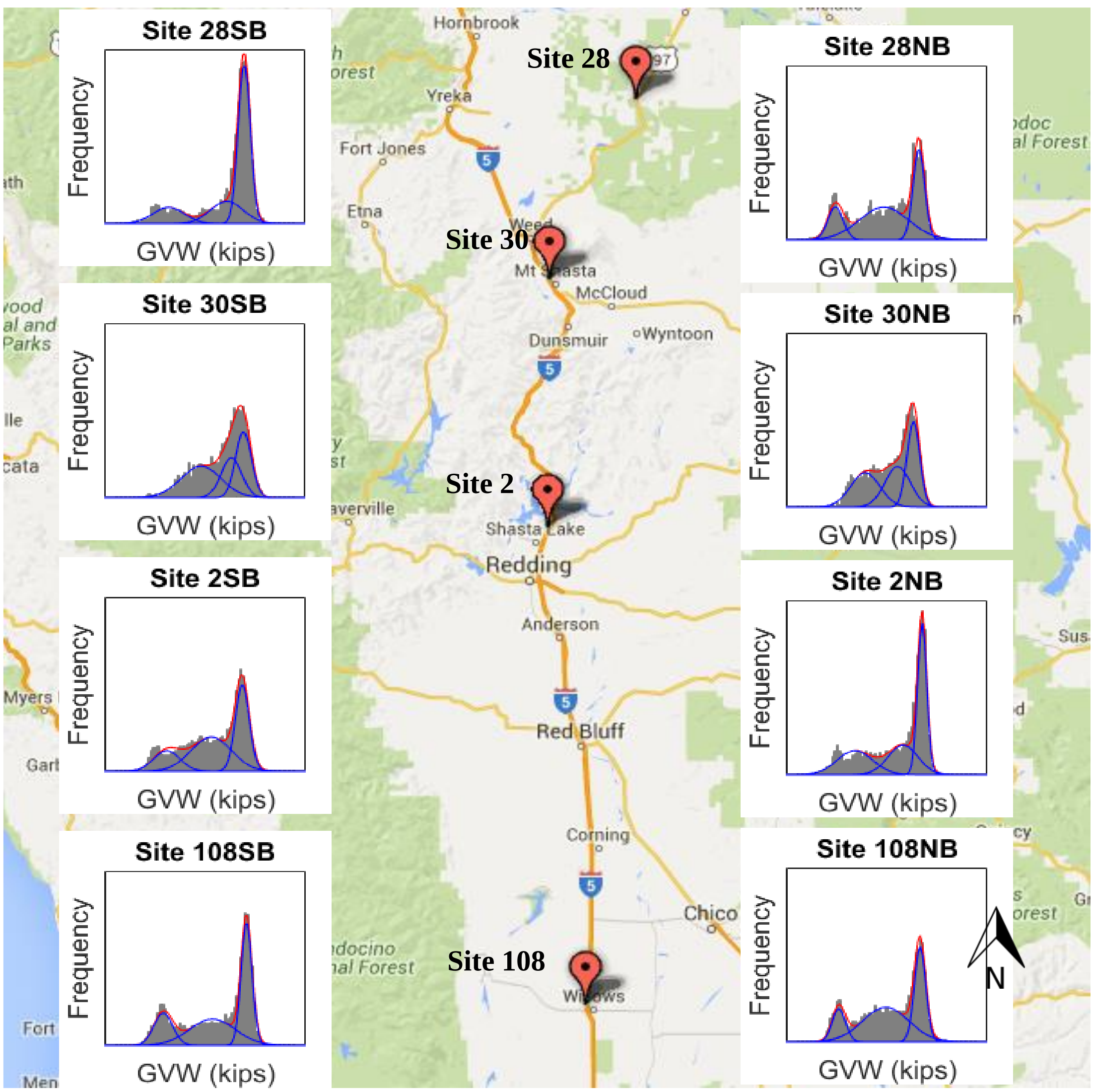


FIGURE 2 Example spatial trends in GVW distributions in northern California.

To combine GVW distributions, a measure of the spatial relationship among sites is needed. Common methods to assess spatial relationships use (a) geographic coordinates to determine Great-Circle Distances (GCD) or (b) shortest path distances between pairs of coordinates. Since GVW distributions depend on direction, regional trends, and route, GCDs and shortest path measures would not fully represent the spatial relationships between sites. For instance, the shortest path between two WIM sites on parallel routes would show the two sites to be highly spatially related when in fact, truck characteristics may differ greatly between the two sites.

GPS data, on the other hand, is better able to reflect these characteristics by providing truck trajectories that capture actual truck travel patterns. Using GPS data, the spatial relation between two sites was defined as the number of truck trajectories that pass between the two sites, referred to as 'shared trips'. Likewise a 'shared site' is a WIM site that has 'shared trips' with the traffic count site. As illustrated in **Figure 3** for traffic count site i, WIM sites A and B are shared sites while site C is not. Site A has three shared trips, Site B has only one shared trip, and Site C has no shared trips with traffic count site i.

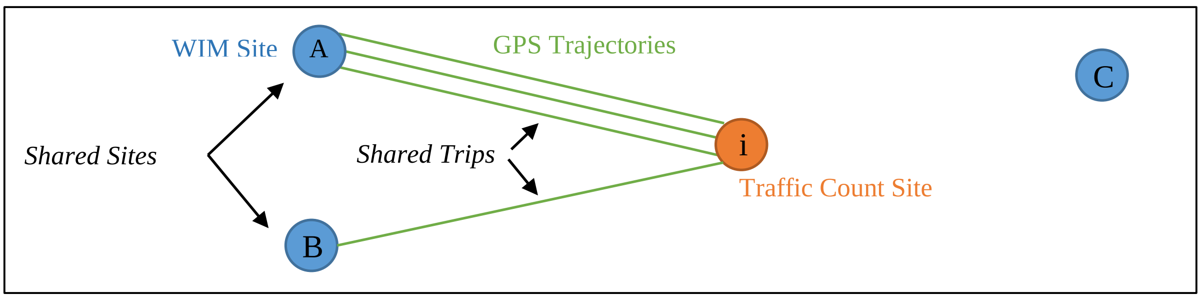


FIGURE 3 Definition of shared sites and shared trips

To address directionality, the number of shared trips would be low for directional sites at the same location (i.e. NB and SB at WIM site i) because it is unlikely that many trucks traverse both directions of a particular station within the same trip. To address regional trends in GVW distributions, sites in the same region that see the same truck traffic would have a higher number of shared truck trips. To address the issue of parallel routes, it is unlikely that the same truck would travel across sensors along parallel routes on the same trip so the number of shared trips between sensors on parallel routes would be low.

Assuming that a higher number of shared trips equates to a higher correlation in the shape of the GVW distributions between the two sites, the number of shared trips can be used as the weighting measure $(W_{i,j})$ to combine GMMs. A higher $(W_{i,j})$ indicates a larger portion of similar truck traffic between two sites and thus a stronger influence on the shape of the GVW distribution while a lower $(W_{i,j})$ indicates a low number of shared trips and thus a weaker influence on the shape of the GVW distribution. A spatial relation matrix was estimated such that each cell contains the number of shared trips, $W_{i,j}$, between traffic count site i and WIM site j. The values of $W_{i,j}$ are used as raw values (i.e. number of shared trips) and do not need to be normalized prior to being input to the GMM. The GMM procedure inherently normalizes mixing proportions to sum to one.

2.4 GVW Distribution Estimation at Traffic Count Sites

Using the defined spatial relation matrix and the estimated GMMs at each WIM site, a GMM was then estimated for each traffic count site. The GMM of the GVW distribution at that site is represented by the equation below. Once the GMM is estimated, the GVW distribution can be found by multiplying $f_j(x)$ by the truck volume recorded at the traffic count site.

$$f_j(x) = \sum_{i=1}^{N} \sum_{m=1}^{M} W_{i,j} \cdot p_{i,m} \cdot \mathcal{N}(x; \mu_{i,m}, \Sigma_{i,m})$$

where
$f_j(x)$ is the GVW weight distribution function for traffic count site *j*
N is the number of WIM sites in the state, i = 1… N
$\mathcal{N}(x; \mu_{i,m}, \Sigma_{i,m})$ is the m^{th} Gaussian distribution for site i with mean $\mu_{i,m}$ and covariance matrix $\Sigma_{i,m}$, m = 1…M
$p_{i,m}$ is the m^{th} Gaussian distribution mixing component for site *i*
$W_{i,j}$ is the spatial measure between traffic count site j and WIM site i extracted from the spatial relation matrix S(i,j)

3 APPLICATION

3.1 Data

The methodology was applied to truck GPS and WIM data from 2010 in California. GPS samples from the American Transportation Research Institute (ATRI) were obtained from four two-week periods in the months of February, May, August and November. The sample of trucks in the ATRI dataset have been shown to be representative of long haul, five axle tractor trailer trucks (*25*). Thus, the GPS dataset is appropriate given that the proposed methodology applies to five axle tractor trailers. There were 112 unidirectional WIM sites reporting data within the time periods corresponding to the GPS dataset as shown in **Figure 4**. To develop the spatial relation matrix, the raw GPS pings were converted to truck trip trajectories. Screenlines were drawn at each of the unidirectional WIM sites to capture the truck trip trajectories passing through each pair of sites. The number of truck trip trajectories that passed through each pair of WIM sites were counted and converted into a spatial relation matrix.

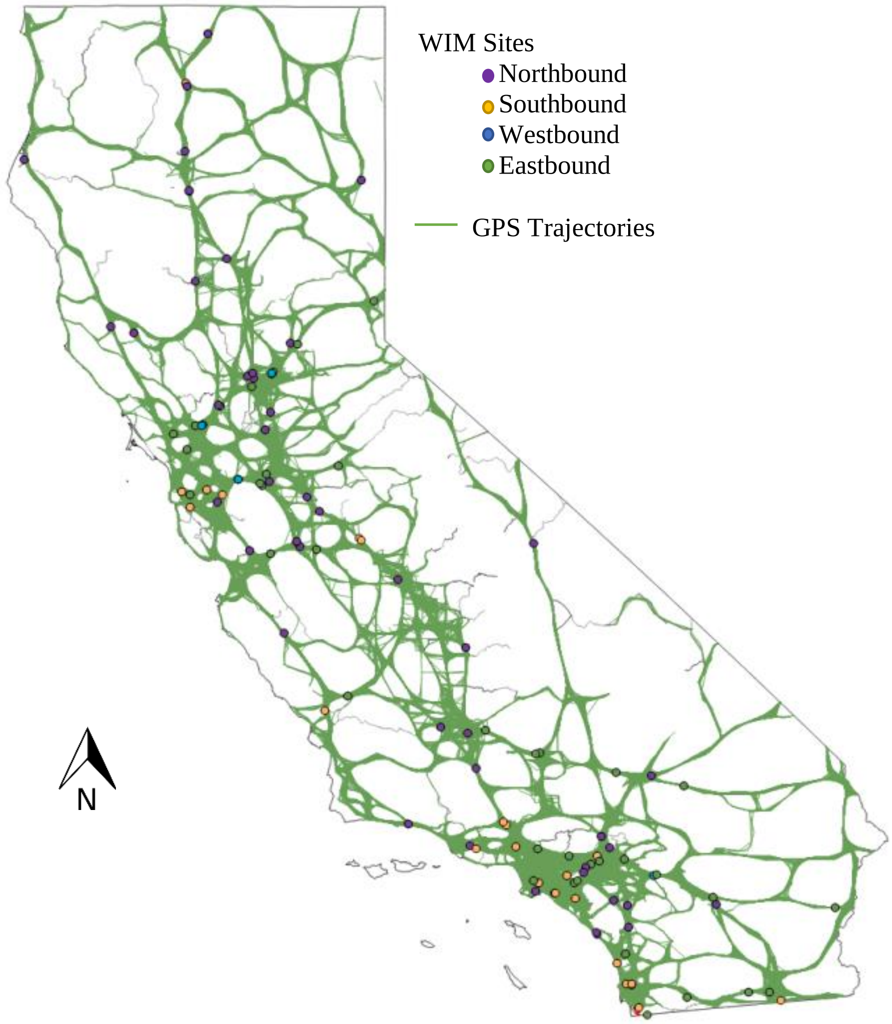


FIGURE 4 GPS truck trip trajectories and WIM station locations.

3.2 Validation

In order to test the accuracy of the proposed method, a leave-one-out cross-validation (LOOCV) approach was adopted. LOOCV entailed treating one WIM site as a traffic count site (i.e. weight distribution unknown) and using the remaining WIM sites to estimate the spatial relation matrix and subsequently estimate the GVW distribution at the WIM site that was held out. This process was repeated for each WIM site in the original sample. To assess the accuracy of the GVW distribution estimation methodology, the observed GVW distribution was then compared to the estimated GVW distribution for each WIM site.

3.3 Results

A requirement of the proposed approach is that the number of shared trips for any traffic count site must be greater than zero otherwise the GMM cannot be estimated. The number of shared trips can be zero if the GPS sample does not show any trips passing over the traffic count site. This can result from an inadequate sample of GPS trajectories or a sparse WIM sensor network. Of the 112 WIM sites, 66 sites had at least one shared site (e.g. a truck trajectory that could be traced from another WIM site). The number of shared sites ranged from one to 22 shared sites with a median of 10 shared sites. The number of shared trips at the 66 sites ranged from one to 2,141 trips with a median value of five trips. The number of mixture components in the estimated GMM at a traffic count site varied by number of shared sites and the number of components in the best-fit GMM (i.e., either two or three) at each of those shared sites. The minimum and maximum number of mixtures were three and 69 mixtures, respectively, with a median of 31 mixtures.

To compare the estimated and observed GVW distributions, coincidence ratios and a two-sample Kolmogorov-Smirnov (KS) statistical test were used under the LOOCV approach. The CR is a relative measure of the fit while the KS test provides a statistical comparison. Coincidence ratios (CR), also called overlapping coefficients, measure the total area that 'coincides' between two distributions and is defined as the ratio of the area in common between two distributions to the total area of the two distributions (*26, 27*). CR takes a value between zero and one with zero indicating disjoint distributions and one indicating completely identical or overlapping distributions. The CR is calculated as follows:

$$CR_i = \frac{\sum_b \min(x_{i,b}, \hat{x}_{i,b})}{\sum_b \max(x_{i,b}, \hat{x}_{i,b})}$$

where
CR_i is the coincidence ratio between the estimated ($\hat{x}_{i,b}$) and observed ($x_{i,b}$) GVW distributions for site i
$x_{i,b}$ is the observed frequency for bin b of the GVW distribution
$\hat{x}_{i,b}$ is the estimated frequency for bin b of the GVW distribution

A two-sample KS test is a nonparametric statistical test for the equality of two continuous distributions. This test is appropriate for trimodal or bimodal distributions as it compares both the location and shape of empirical and estimated distributions. The hypotheses for the two sample KS test are:

H_o: the estimated distribution is from the same continuous distribution as the measured distribution

H_a: the estimated distribution is from a different continuous distribution than the measured distribution

A histogram and cumulative distribution of the CR are shown in **Figure 5**. Of the 66 sites with estimated GVW distributions, 26 (40%) had CR greater than 0.70 and 46 (70%) had statistically similar distributions according to the KS test at the 1% significance level.

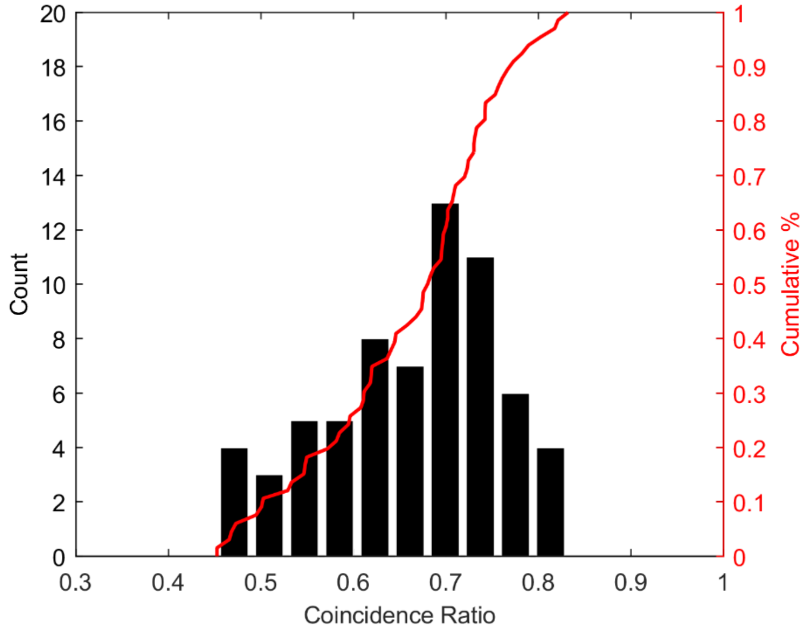


FIGURE 5 PDF and CDF of coincidence ratios.

Figure 6 compares the measured and estimated GVW distributions for a select set of WIM sites. The denomalization procedure of multiplying normalized GVW estimates by 10 kips, the assumed steering axle weight, was applied prior to plotting the given examples. The top two examples (**Figure 6a and b**) have the highest CR values, the center two examples (**Figure 6c and d**) have CR corresponding to the median CR of 0.66, and the bottom two examples (**Figure 6e and f**) have the lowest CR. In general, the estimated GMM with high CR match the GVW range and location of peaks shown in the measured GVW distribution. Estimated GMM with CR in the median range tend not to match the height of peaks of the measured data as in **Figure 6c**. Estimated GMM with lower CR tend to have shifted GVW ranges as in **Figure 6e** and low traffic volumes as in **Figure 6f**. Shifts in the GVW range are likely indications of WIM sensor calibration issues that were not corrected in the normalization procedure. **Figure 6f** is a case of a site with relatively low traffic volumes which led to a poor fit GMM although the general shape of the distribution is mimicked by the estimated GMM.

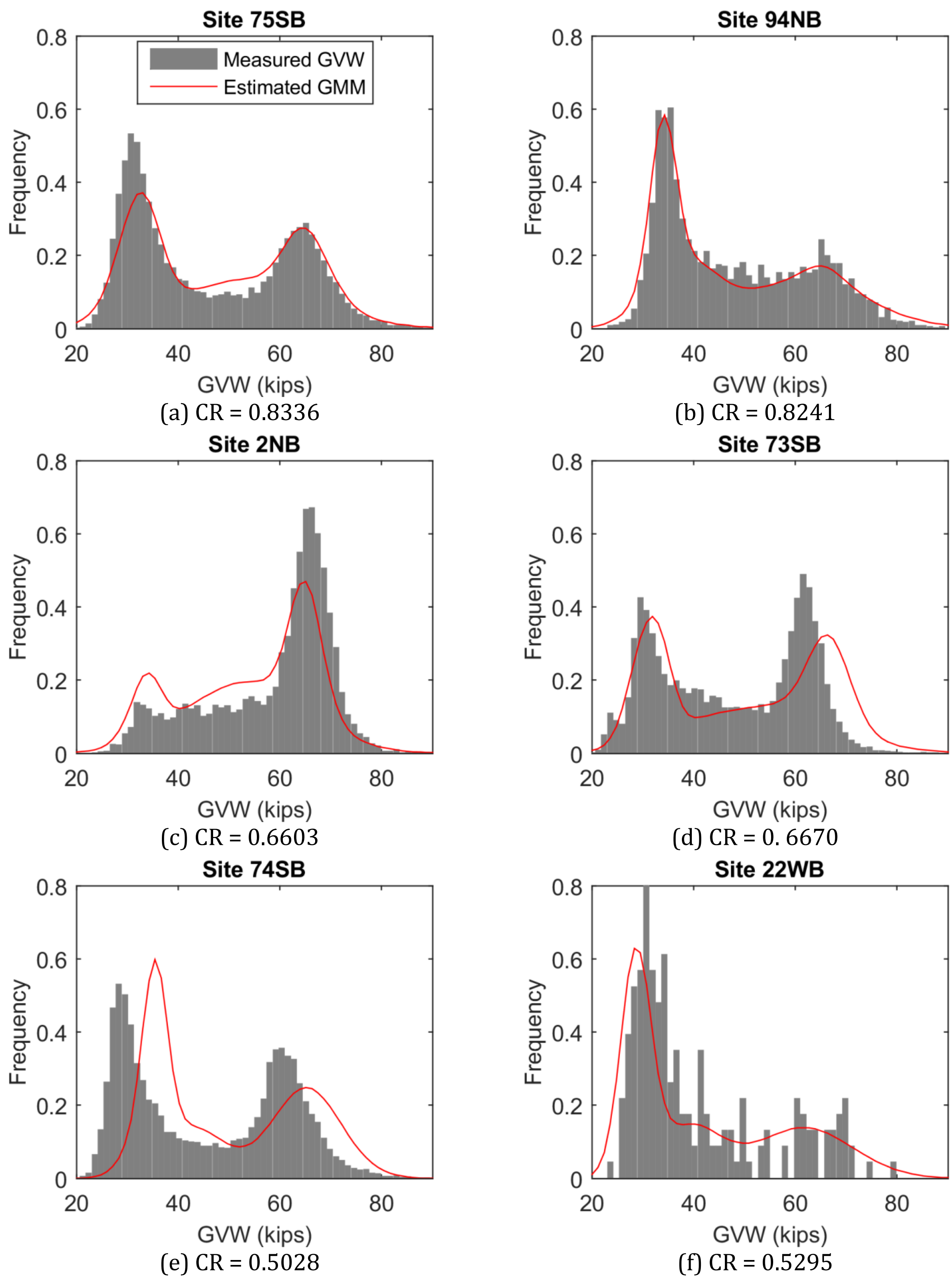


FIGURE 6 Estimated GVW distributions for select sites.

3.4 Sensitivity Analysis

The number of mixtures, shared trips, and shared sites have an effect on the accuracy of the estimated GMMs. These measures are controlled by the assumed spatial relationships between sites which in turn determine which sites to include in the GMM. To test the robustness of the methodology to the definition of spatial relationships, the following sensitivity analyses were performed:

(**A1**) *Threshold on the number of shared trips:* The threshold on the number of shared trips restricts a site from being included in the GMM unless the number of shared trips between the traffic count and the WIM is above the threshold. Thresholds of 10, 100, and 200 shared trips were compared.

(**A2**) *Distance-weighted spatial relation matrix:* The spatial weight matrix contains the number of shared trips ($W_{i,j}$) between pairs of stations. The matrix was weighted by the inverse of the distance ($1/d_{ij}$) between pairs of sites. The inverse of distance allows the spatial relation to decrease as the stations become farther apart. GCD measures the shortest distance between two points 'as the crow flies' and was use to define distance.

(**A3**) *GCD-based spatial relation matrix:* The GCD-based approach represents what could be done without GPS data by using only the inverse of the GCD between sites to define the spatial relation matrix.

To determine whether the alternate definition of spatial relationships improved model performance, (A1), (A2), and (A3) were assessed by comparing each site's CR result. For each site (i), the change in the CR was calculated as: $\Delta \mathrm{CR}_\mathrm{i} = \mathrm{CR}_{\mathrm{proposed,i}} - \mathrm{CR}_{\mathrm{alternate,i}}$ where $\mathrm{CR}_{\mathrm{proposed,i}}$ is the CR of the proposed method for site i and $\mathrm{CR}_{\mathrm{alternate,i}}$ is the CR of the alternate model for site i. **Figure 7** illustrates the model performance and number of sites estimated for A1 to A3 compared to the proposed model.

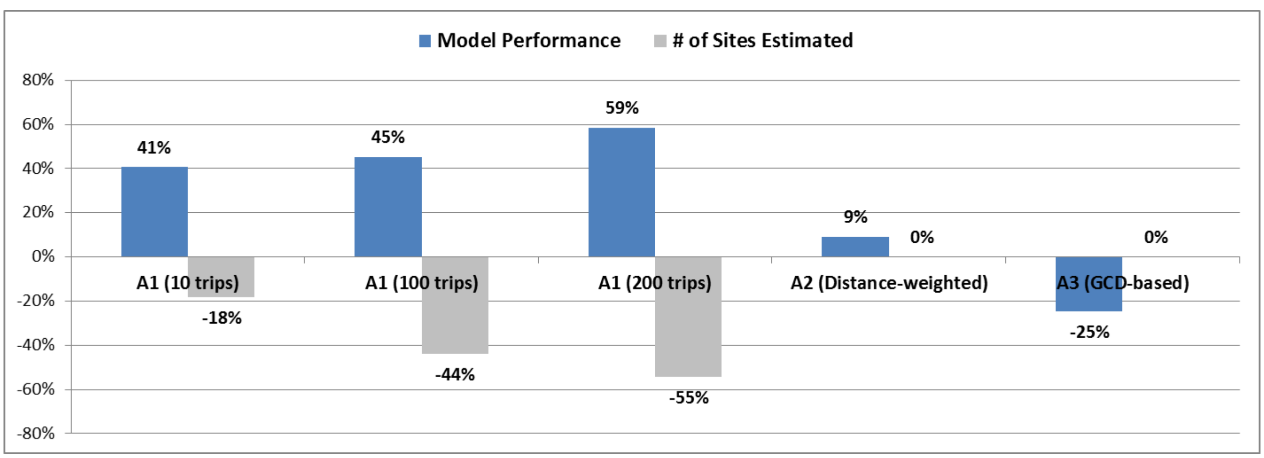
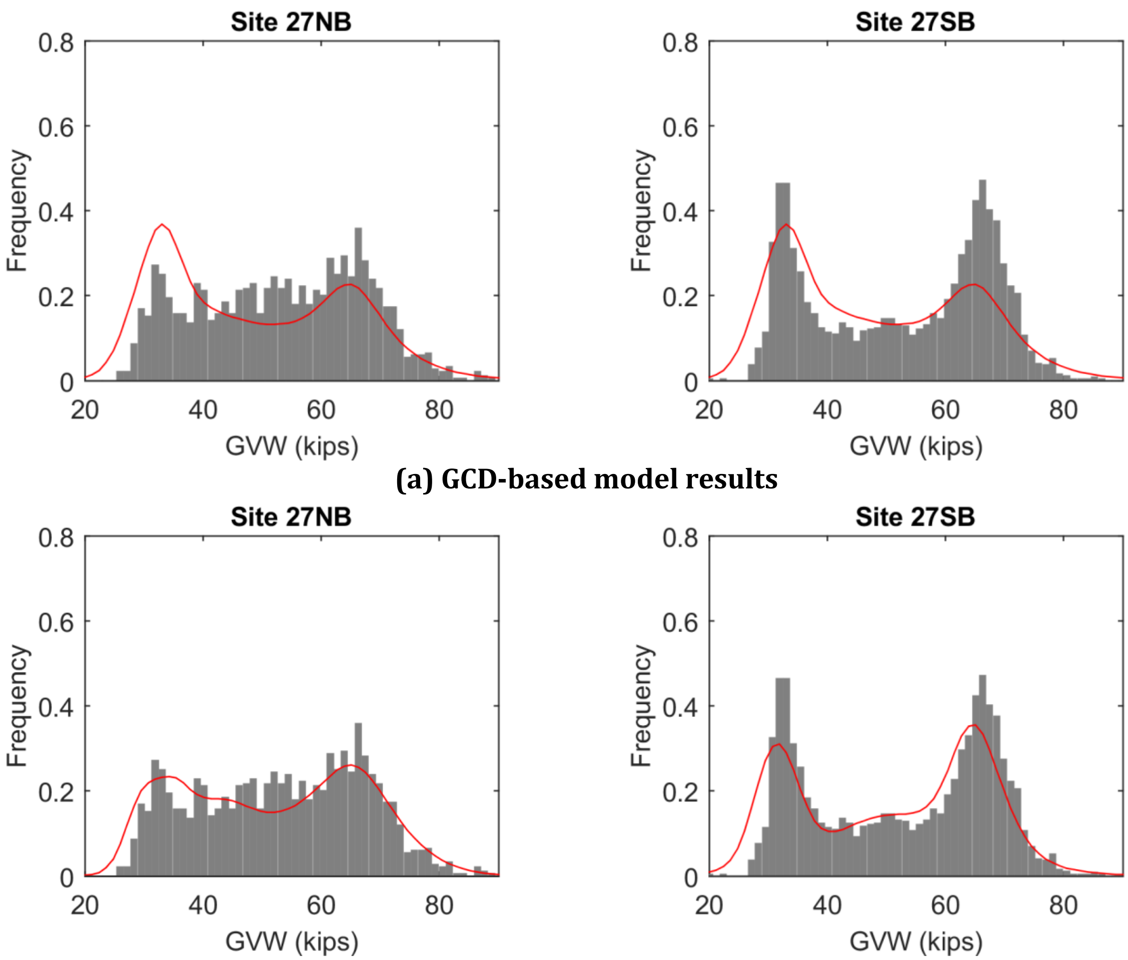


FIGURE 7 Results of the sensitivity analysis.

Based on the results of **Figure 7**, thresholding the shared trips at 10, 100 and 200 trips reduced the number of sites by 18, 44, and 55% respectively, compared to the proposed model but significantly improved the performance for the estimated sites. For example, 41% of estimated sites showed improved model performance in the A1 model with a threshold of 10 shared trips. The tradeoff in the number of sites estimated and performance must be addressed when applying the proposed method. If one desired to estimate truck weights at as many sites as possible then any number of shared trips is appropriate, however if more emphasis is placed on performance, then thresholds should be applied. The distance-weighted method (A2) improved the model performance to 9% without sacrificing the number of estimated sites.

The GCD-based model (A3) reduced model performance for 25% of sites compared to the proposed model. The GCD-based approach produces GMMs that are averages of GVWs within the region rather than specific estimates based on truck travel patterns. This approach is not able to capture unique peaks representing relatively higher volumes of empty or loaded trucks that might be seen at a particular site. This is most evident for directional stations which tend to exhibit different GVW distributions but through the distance-based approach have identical GMMs. Note that the GCD between sites of opposing directions at the same station (e.g. NB and SB at site i) would be zero and that the distance matrix is symmetrical. **Figure 8** contrasts the proposed method to the distance-based approach for the NB and SB directions and highlights the ability of the proposed method to produce GVW distributions that differ by direction for the same site.



(a) GCD-based model results

(b) Proposed model results

FIGURE 8 Comparisons of measured and estimated GVW distributions for the GCD-based and proposed methods.

4 CONCLUSIONS

This paper presents a method to estimate GVW distributions of '3-S2' trucks at traffic count sites based on WIM and GPS data. The purpose is to increase the spatial resolution of truck weight data using existing data collection platforms. Comparisons of GVW distributions estimated by the proposed method to the measured GVW data show favorable results for the 112 WIM sites in California. The sensitivity analysis demonstrated that the proposed method outperformed a baseline approach that relies only on geographic distances between sites.

Several enhancements to the methodology can be carried out to improve the results. First, a more advanced normalization approach such as that by Chou and Nichols (*20*) can be applied to the GVW data at each site prior to determining GMM parameters. This would reduce the effects of shifted GVW ranges evidenced in the estimated GVW distributions. Second, the methodology presented concentrates on GVW of '3-S2' trucks and could be expanded for other common truck configurations or to determine axle load spectra in addition to GVW. Third, a sensitivity analysis to examine the effect of WIM site density can be undertaken to determine the potential of other states to apply the proposed method. Similarly, the spatial and industry coverage of the GPS sample plays a role in the accuracy of the method presented in this paper. Thus, a sensitivity analysis examining the effect of GPS sample characteristics is warranted. Fourth, a shortest path distance based on truck routes between sites can be explored as a potential spatial relation measure for the proposed method. This would require a shortest path to be found between every traffic count and WIM site that considers truck route restrictions. While not impossible, this would be a sizeable problem for dense road networks. Lastly, GPS is only one of several means of gathering truck trajectory data. Tracking capable sensors such as smartphones, truck weight station pre-clearance programs such as PrePass (*28*), and connected vehicle systems are all viable substitutes for the GPS data used in the methodology presented.

Truck weight data is a much needed resource for pavement and bridge design and performance, traffic monitoring, freight transportation planning, and emissions estimation. The method in this paper allows truck weight data to be estimated at a larger number of sites, providing input data for a number of diverse applications. Moreover, the methodology could be used to locate new WIM sites with the goal of maximizing the number of traffic count sites for which a GVW distribution can be estimated.

5 REFERENCES

1. Hall, K., Xiao, X., and Wang, C.P., "Calibration of the M-E Design Guide", Final Report – TRC-1003, Arkansas State Highway and Transportation Department, Little Rock, Arkansas, November 2014.

2. Tran, N.H., and K.D. Hall. "Development and Influence of Statewide Axle Load Spectra on Flexible Pavement Performance," Transportation Research Record No. 2037, Transportation Research Board, Washington, DC, 2007.

3. Battelle, FAF^3 Freight Traffic Analysis, March 2011.

4. Jeong, K., Tok, A., and Ritchie, S.G., California Vehicle Inventory and Use Survey: Pilot Study Insights, Presented at the Annual Meeting of the Transportation Research Board, Washington, D.C., 2016.

5. California Air Resources Board (CARB), 2016. EMFAC2011-LDV User's Guide. [pdf] Available at: <http://www.arb.ca.gov/msei/emfac2011-ldv-users-guidefinal.pdf> (accessed March 3, 2016).

6. Federal Highway Administration, LTBP Program's Literature Review on Weigh-in-Motion Systems, FHWA Publication No. FHWA-HRT-16-024, June 2016.

7. California Department of Transportation (Caltrans), California Public Road Data: Statistical Information Derived from the Highway performance Monitoring System. [pdf] Available at: <http://www.dot.ca.gov/hq/tsip/hpms/hpmslibrary/prd/2013prd/2013PublicRoadData.pdf> (accessed June 2016).

8. Quinley, R. WIM Data Analysts' Manual, Report No. FHWA-IF-10-018, Federal Highway Administration, Washington, D.C. 2010.

9. Traffic Monitoring Guide. Publication FHWA-PL-01-021. FHWA, U.S. Department of Transportation, Washington, D.C., 2001.

10. California Department of Transportation (Caltrans), Performance Measurement System (PeMS), Available at <http://pems.dot.ca.gov/> (accessed June 2016)

11. Jeng, S.T., Chu, L. and Hernandez, S., Wavelet-k nearest neighbor vehicle classification approach with inductive loop signatures. Transportation Research Record: Journal of the Transportation Research Board, No. 2380, pp.72-80, 2013.

12. Hernandez, S., Tok, A., and Ritchie, S.G., Integration of Weigh-in-Motion (WIM) and inductive signature data for truck body classification, Transportation Research Part C: Emerging Technologies, Vol. 68, pp. 1-21, 2016.

13. Nichols, A., and D. M. Bullock, 2004. Quality Control Procedures for Weigh-in-Motion Data. Publication FHWA/IN/JTRP-2004/12. Joint Transportation Research Program, Indiana Department of Transportation and Purdue University, West Lafayette, Indiana.

14. Papagiannakis, A.T., Quinley, R., Brandt, S.R. NCHRP Synthesis Project 386: High Speed Weigh-in-Motion System Calibration Practice. Transportation Research Board of National Academies, Washington, D.C., 2008.

15. Prozzi, J.A. and Hong, F., Effect of Weigh-in-Motion System Measurement Errors on Load-Pavement Impact Estimation, Journal of Transportation Engineering, Vol. 133, No. 1, 2007, pp 1-10.

16. Dahlin, C. Proposed Method for Calibrating Weigh-in-Motion Systems and for Monitoring That Calibration Over Time. In Transportation Research Record 1364, TRB, National Research Council, Washington D.C., 1992, pp. 161–168.

17. Han, C., W. T. Boyd, and M. M. Marti. Quality Control of Weigh-in- Motion Systems Using Statistical Process Control. In Transportation Research Record 1501, TRB, National Research Council, Washington, D.C., 1995, pp. 72–80.

18. Southgate, H.F., Quality Assurance of Weigh-In-Motion Data, Washington, D.C., Federal Highway Administration, <https://www.fhwa.dot.gov/ohim/tvtw/wim.pdf> (Access October 2014)

19. Jeng, S.T., Chu, L., and Cetin, M., Weigh-In-Motion Station Monitoring and Calibration Using Inductive Loop Signature Technology, Presented at the 95th Annual Meeting of the Transportation Research Board, Washington, D.C., 2015.

20. Chou, C.S. and Nichols, A., A Recommended Procedure to Adjust Inaccurate Weigh-in-Motion Data, Presented at the 95th Annual Meeting of the Transportation Research Board, Washington, D.C., 2015.

21. Hastie, T., Tibshirani, R., Friedman, J., The elements of statistical learning, New York: Springer, Vol. 2, No. 1, 2009.

22. Nichols A. and Cetin, M., Numerical Characterization of Gross Vehicle Weight Distributions from Weigh-in-Motion Data, Transportation Research Record: Journal of the Transportation Research Board, No. 1993, Transportation Research Board of the National Academies, pp. 148-154, 2007.

23. Hyun, K., Hernandez, S., Tok, A., and Ritchie, S.G., Truck Body Configuration Volume and Weight Distribution: Estimation by Using Weigh-in-Motion, Transportation Research Record: Journal of the Transportation Research Board, No. 2478, Transportation Research Board of the National Academies, pp. 103-112, 2015.

24. Akaike, H. Likelihood of a Model and Information Criteria. Journal of Econometrics, Vol. 16, 1981, pp. 3–14.

25. Zanjani, A.B., Pinjari, A.R., Kamali, M., Thakur, A., Short, J., and Tabatabaee, S., Estimation of Statewide Origin-Destination Truck Flows Using Large Streams of GPS Data: An Application for the Florida Statewide Model, Annual Meeting of the Transportation Research Board, Washington, D.C., January, 2015.

26. Federal Highway Administration (FHWA), Quick Response Freight Manual II, accessed online April 22, 2016 at http://ops.fhwa.dot.gov/freight/publications/qrfm2/index.htm, 2007.

27. Henry F. Inman; Edwin L. Bradley Jr (1989). The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities. Communications in Statistics - Theory and Methods, 18(10), 3851-3874.

28. PrePass, Available online at <http://prepass.com/>, (accessed June 2016).